

RECEIVED

JUN 12 2002



TECH CENTER 1600/2900

1652

## RAW SEQUENCE LISTING

DATE: 05/31/2002

PATENT APPLICATION: US/09/699,136

TIME: 08:35:11

Input Set : D:\30062-20049.txt

Output Set: N:\CRF3\05312002\I699136.raw

ENTERED

4 <110> APPLICANT: Kosan Biosciences, Inc.  
 5 Santi, Daniel  
 6 Peck, Larry  
 7 Dayem, Linda  
 8 Kealey, James  
 10 <120> TITLE OF INVENTION: HETEROLOGOUS PRODUCTION OF POLYKETIDES  
 13 <130> FILE REFERENCE: 30062-20049.00  
 15 <140> CURRENT APPLICATION NUMBER: US 09/699,136  
 16 <141> CURRENT FILING DATE: 2000-10-27  
 18 <150> PRIOR APPLICATION NUMBER: US 60/161,414  
 19 <151> PRIOR FILING DATE: 1999-10-25  
 21 <150> PRIOR APPLICATION NUMBER: US 60/161,703  
 22 <151> PRIOR FILING DATE: 1999-10-27  
 24 <150> PRIOR APPLICATION NUMBER: US 60/206,082  
 25 <151> PRIOR FILING DATE: 2000-05-18  
 27 <160> NUMBER OF SEQ ID NOS: 2  
 29 <170> SOFTWARE: FastSEQ for Windows Version 4.0  
 31 <210> SEQ ID NO: 1  
 32 <211> LENGTH: 447  
 33 <212> TYPE: DNA  
 34 <213> ORGANISM: Artificial Sequence  
 36 <220> FEATURE:  
 37 <223> OTHER INFORMATION: Isolated and recombinant form of the full  
 38 epimerase gene sequence

W--&gt; 40 &lt;221&gt; NAME/KEY: CDS

41 &lt;222&gt; LOCATION: (1)...(444)

W--&gt; 43 &lt;400&gt; 1

44	atg agt aat gag gat ctt ttc atc tgt atc gat cac gtg gca tat gcg	48
45	Met Ser Asn Glu Asp Leu Phe Ile Cys Ile Asp His Val Ala Tyr Ala	
46	1 5 10 15	
48	tgc ccc gac gcc gac gag gct tcc aag tac tac cag gag acc ttc ggc	96
49	Cys Pro Asp Ala Asp Glu Ala Ser Lys Tyr Tyr Gln Glu Thr Phe Gly	
50	20 25 30	
52	tgg cat gag ctc cac cgc gag gag aac ccg gag cag gga gtc gtc gag	144
53	Trp His Glu Leu His Arg Glu Glu Asn Pro Glu Gln Gly Val Val Glu	
54	35 40 45	
56	atc atg atg gcc ccg gct gcg aag ctg acc gag cac atg acc cag gtt	192
57	Ile Met Met Ala Pro Ala Ala Lys Leu Thr Glu His Met Thr Gln Val	
58	50 55 60	
60	cag gtc atg gcc ccg ctc aac gac gag tcg acc gtt gcc aag tgg ctt	240
61	Gln Val Met Ala Pro Leu Asn Asp Glu Ser Thr Val Ala Lys Trp Leu	
62	65 70 75 80	
64	gcc aag cac aat ggt cgc gcc gga ctg cac cac atg gca tgg cgt gtc	288

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/699,136

DATE: 05/31/2002

TIME: 08:35:11

Input Set : D:\30062-20049.txt

Output Set: N:\CRF3\05312002\I699136.raw

```

65 Ala Lys His Asn Gly Arg Ala Gly Leu His His Met Ala Trp Arg Val
66          85          90          95
68 gat gac atc gac gcc gtc agc gcc acc ctg cgc gag cgc ggc gtg cag      336
69 Asp Asp Ile Asp Ala Val Ser Ala Thr Leu Arg Glu Arg Gly Val Gln
70          100          105          110
72 ctg ctg tat gac gag ccc aag ctc ggc acc ggc ggc aac cgc atc aac      384
73 Leu Leu Tyr Asp Glu Pro Lys Leu Gly Thr Gly Gly Asn Arg Ile Asn
74          115          120          125
76 ttc atg cat ccc aag tcg ggc aag ggc gtg ctc atc gag ctc acc cag      432
77 Phe Met His Pro Lys Ser Gly Lys Gly Val Leu Ile Glu Leu Thr Gln
78          130          135          140
80 tac ccg aag aac tga      447
81 Tyr Pro Lys Asn
82 145
85 <210> SEQ ID NO: 2
86 <211> LENGTH: 148
87 <212> TYPE: PRT
88 <213> ORGANISM: Artificial Sequence
90 <220> FEATURE:
91 <223> OTHER INFORMATION: Deduced amino acid sequence of the epimerase gene
92     sequence
94 <400> SEQUENCE: 2
95 Met Ser Asn Glu Asp Leu Phe Ile Cys Ile Asp His Val Ala Tyr Ala
96  1          5          10          15
97 Cys Pro Asp Ala Asp Glu Ala Ser Lys Tyr Tyr Gln Glu Thr Phe Gly
98          20          25          30
99 Trp His Glu Leu His Arg Glu Glu Asn Pro Glu Gln Gly Val Val Glu
100          35          40          45
101 Ile Met Met Ala Pro Ala Ala Lys Leu Thr Glu His Met Thr Gln Val
102          50          55          60
103 Gln Val Met Ala Pro Leu Asn Asp Glu Ser Thr Val Ala Lys Trp Leu
104 65          70          75          80
105 Ala Lys His Asn Gly Arg Ala Gly Leu His His Met Ala Trp Arg Val
106          85          90          95
107 Asp Asp Ile Asp Ala Val Ser Ala Thr Leu Arg Glu Arg Gly Val Gln
108          100          105          110
109 Leu Leu Tyr Asp Glu Pro Lys Leu Gly Thr Gly Gly Asn Arg Ile Asn
110          115          120          125
111 Phe Met His Pro Lys Ser Gly Lys Gly Val Leu Ile Glu Leu Thr Gln
112          130          135          140
113 Tyr Pro Lys Asn
114 145

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/699,136

DATE: 05/31/2002

TIME: 08:35:12

Input Set : D:\30062-20049.txt

Output Set: N:\CRF3\05312002\I699136.raw

L:40 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!  
L:43 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:1